The invention herein pertains to comparison between a recognized sequence of labels and a stored sequence of labels. An input signal is received, and the input signal is compared with stored label models in order to generate the recognized sequence of labels. In addition, confidence data is also generated, representative of the confidence that the recognized sequence of labels is representative of the input signal. A measure of similarity is obtained by comparing the recognized sequence of labels with a stored sequence of labels by using a combination of two pieces of information: predetermined confusion data which defines confusability between different labels, and the generated confidence data.

In this regard, Applicants wish to distance themselves with remarks previously submitted that might be interpreted to imply that the claimed invention only uses confidence data for an entire input sequence. Any such implication is incorrect; although it is certainly possible that embodiments of the invention might do so, it is not strictly necessary given the language of the claims and given the breadth of the invention. In particular, it is clear from the embodiments of the invention (and also from many of the dependent claims such as Claim 3 herein) that the invention also covers a system which uses confidence data associated with each label within the sequence. Accordingly, any such comments to the contrary, such as those that might be found in the Supplemental Response dated December 3, 2004, are hereby withdrawn explicitly, and there should be no reliance on any such comments either in this application or any patent issuing herefrom.

It is therefore a feature of the invention that a measure of similarity is obtained not only on confidence data generated during recognition of an input signal, but

also based on predetermined confusion data which defines confusability between different labels. By virtue of this feature, and since the measure of similarity is obtained also based on confusability between the different labels against which the input signal was recognition-processed, the resulting similarity measure can provide a more accurate and more realistic measure of the similarity between the recognized sequence of labels and the stored sequence of labels. In effect, since it is always possible to encounter recognition errors because of mis-recognition of the input signal, particular as regards confusability between the different labels themselves, the similarity measure of the present invention at least allows more robust detection and reduction of the problem.

Although the Office Action indicated allowable subject matter, it also entered a rejection of all independent claims under 35 U.S.C. § 102(b) over U.S. Patent 5,737,489 (Chou). Rejections of other claims were also entered, over Chou or Chou in various combinations with U.S. Patent 6,662,180 (Aref) and U.S. Patent 5,333,275 (Wheatley). All these rejections are respectfully traversed.

Chou describes a verification system for verifying recognition results of a speech recognition processor. As shown in Figure 1, unknown speech 18 is fed to recognition processor 10, resulting in a hypothesized recognition string signal 20. String 20 is subjected to verification results in verification processor 14 which outputs a confidence measure signal 22. If the confidence exceeds a specific threshold, the threshold comparator 24 outputs a verification decision indicative of whether the recognition processor has or has not accurately recognized the unknown speech.

More details of verification processor 14 are shown in Chou's Figure 2. In this regard, because of poor patent draftsmanship by the draftsman of the Chou patent, reference numerals from Figure 1 are not accurately repeated in Figure 2. Nevertheless, it is clear that all of Figure 2, with the exception of threshold comparator 40 (which corresponds to threshold comparator 24 in Figure 1), constitute verification processor 14 of Figure 1. It is further clear that Figure 2's label of "hypothesis speech 55" corresponds to Figure 1's "hypothesis recognized string signal 20". With this understanding in mind, operation of Chou's verification processor 14, and his overall recognition system, becomes more clear.

In particular, it is clear that the equations in Chou's column 8 do not use confidence data; rather, they generate it. These equations discuss operation of the verification processor 14, and it is the verification processor that generates the confidence signal 22. Put another way, in Chou's system, recognition processor 10 generates a hypothesis of possible recognition results. These results do not yet have any quality associated with them, and they may or may not be correct. It is the purpose of verification processor 14 to generate some measure of confidence of these recognition results, as signified at reference numeral 22 in Figure 1. The confidence signal is compared against a threshold to determine the accuracy of the actual recognition made by processor 10. See Chou, column 4, lines 34 to 51:

> "The recognition processor receives as input an unknown speech string 18 (an utterance) of words. The recognition processor 10 accesses the recognition database 12 in response to the unknown speech string 18 input and scores the unknown speech string of words against the recognition models in the recognition database 12 to classify the unknown

string of words and to generate a hypothesis recognized string signal. The verification processor receives the hypothesis string signal 20 as input to be verified. The verification processor 14 accesses the verification database 16 to test the hypothesis string signal against verification models stored in the verification database. Based on the verification test, the verification processor 14 generates a confidence measure signal 22. The confidence measure signal is passed to a threshold comparator 24 to be compared against a verification threshold signal value to determine the accuracy of the classification decision made by the recognition processor 10."

It is therefore incorrect to conclude that the equations shown in Chou's column 8 rely on "confidence data representative of the confidence that a recognized sequence of labels is representative of an input signal", as set out in the claims herein, for the simple reason that the entire purpose of Chou's equations is to generate such confidence data. Applicants wish to be clear on this point, since they are not taking the position that Chou does not generate confidence data. Rather, it is the position of Applicants that although Chou generates confidence data, it does not use such confidence data to obtain a measure of similarity by use of both the confidence data and predetermined confusion data which defines confusability between different labels.

Moreover, even if the Office Action's interpretation of Chou's equations were accepted as correct, which is not conceded as described above, the Office Action's attempt to equate various elements of these equations with features in the claims is faulty.

First, as corrected pointed out by the Examiner, the component $g_i(O_q)$ corresponds to Chou's word model score. This score is obtained by the verification processor 14 by comparing a sequence of feature vectors output bu the feature extractor 28 with an HMM-based word model. However, such a word model score cannot be

considered to correspond to the claimed confidence data as it is possible to have a high word model score and a low confidence, or to have a low word model score and a high confidence. For example, if the user utters the word "there", the word model score for the models "there", "their" and "they're" will all have high word model scores. However, because of the alternatives, the confidence associated with the recognized word model will be low. In contrast, where a relatively long word is spoken which does not match well with its corresponding word model (resulting in a low word model score), a high confidence can still be obtained if there are no alternative hypotheses for the uttered word. Therefore, it is wrong to compare the word model score component of equation 2 with the claimed confidence data.

Second, with regard to the claimed confusion data, Claim 1 specifies that it is "predetermined" whereas any confusion data of Chou is not predetermined and is calculated differently for each word being verified. This is rooted in the further difference that the claimed confusion data defines confusability between different labels, whereas in Chou, confusion is not between labels themselves (or, using Chou's terminology, the "anti-keywords") but rather is confusion between the recognized sequence and the anti-keywords. More specifically, the Office Action states that the term $G_l(O_q)$ corresponds to the claimed confusion data. However, as is clear from column 8 lines 33 to 40 and from equation 3, the confusion data of Chou is obtained by comparing the input speech signal (O_q) with anti-keyword HMMs ($\theta_l^{(a)}$) and an acoustic filler model HMM($\theta^{(f)}$). It is therefore clear that the component $G_l(O_q)$ given in equation 2 of Chou cannot be the

confusion data of Claim 1, particularly since the confusion data defined by Claim 1 is both "predetermined" and "between different labels".

It is therefore respectfully submitted that the claims herein define subject matter that is neither anticipated nor would have ben obvious from Chou, or from Chou in any permissible combination with Aref or Wheatley. Withdrawal of the rejections is respectfully requested.

REQUEST FOR INTERVIEW

It is respectfully requested that the Examiner contact the undersigned attorney at (714) 540-8700 when the case is next taken up for action, so as to schedule an interview.

Applicants' undersigned attorney may be reached in our Costa Mesa, California office at (714) 540-8700. All correspondence should continue to be directed to our below-listed address.

Respectfully submitted,

Michael K. O'Neill

Attorney for Applicants
Michael K. O'Neill
Registration No.: 32,622

FITZPATRICK, CELLA, HARPER & SCINTO
30 Rockefeller Plaza
New York, New York 10112-2200
Facsimile: (212) 218-2200

CA_MAIN 99392v1